Active learning and optimal climate policy

Article (Accepted Version)

Active learning and optimal climate policy

Abstract

This paper develops a climate-economy model with uncertainty, irreversibility, and active learning. Whereas previous papers assume learning from one observation per period, or experiment with control variables to gain additional information, this paper considers active learning from investment in monitoring, specifically in improved observations of the global mean temperature. We find that the decision maker invests a significant amount of money in climate research, far more than the current level, in order to increase the rate of learning about climate change. This helps the decision maker make improved decisions. The level of uncertainty decreases more rapidly in the active learning model than in the passive learning model with only temperature observations. As the uncertainty about climate change is smaller, active learning reduces the optimal carbon tax. The greater the risk, the larger is the effect of learning. The method proposed here is applicable to any dynamic control problem where the quality of monitoring is a choice variable, for instance, the precision at which we observe GDP, unemployment, or the quality of education.

Key words

Climate policy; irreversibility; uncertainty; learning; active learning

JEL Classification

Q54; O3; C63

1 Introduction

It has long been known that expectations on future information affect current decisions. This is typically specified as *passive learning*, where the decision maker updates information about state variables as model time progresses. In some cases, it might be optimal to *learn by experimenting* with control variables. In this paper, we consider *active learning*, i.e., investment in improved monitoring, as an alternative route to increase information, and investigate the implications for climate policy.

There is considerable uncertainty about all aspects of climate change and climate policy. As the consequences of emission abatement decisions are irreversible or at least long-lived, the prospect of future learning, materially affects optimal climate policy in the short-run. The previous literature assumes that new knowledge either arrives exogenously or arrives without cost in the form of new observations about the climate system.[1] However, learning requires investment and can be decelerated or accelerated by monitoring and research. In this paper, we explore simultaneous decisions about learning and abatement.

Uncertainty becomes smaller as the decision maker observes the system. As the state of the system, and thus the information it contains, depends on past decisions, the planner controls, to a certain extent, what is learned. However, in the literature (e.g., Kelly and Kolstad, 1999; Kelly and Tan, 2015; Hwang et al., 2017), learning is a by-product of decisions on the carbon tax. In such a representation, there are two state variables (e.g., the carbon stock and the stock of knowledge) but only one control variable (the carbon tax). Decisions

[1] Manne and Richels (1992), Peck and Teisberg (1993), Kolstad (1996a, b), Nordhaus and Popp (1997), Ulph and Ulph (1997), Gollier et al. (2000), and Webster (2002) incorporate exogenous learning. In these studies, information arrives exogenously at some points in time, and thus learning is independent of actions of the decision maker. Kelly and Kolstad (1999), Leach (2007), Webster et al. (2008), Cai et al., (2012), Jensen and Traeger (2013), Lemoine and Traeger (2014), Kelly and Tan (2015), Fitzpatrik and Kelly (2015), Rudik (2016), and Hwang et al. (2017) construct endogenous learning models. The cost of learning is not explicitly considered in their papers, and learning happens as new instances of state variables are realized. Karp and Zhang (2006), Van Wijnbergen and Willem (2015) investigate the effect of such passive learning with theoretical models.

are therefore second best optimal (Tinbergen, 1954). We introduce a second control variable (i.e. research investment) in order to control learning separately from emissions.

In a seminal paper, Prescott (1972) shows that there is a tradeoff between control and information when there is uncertainty about the effect of a policy instrument. He finds it is optimal to sacrifice part of the current benefits of control in order to obtain information that improves future decisions – we refer to this as (*learning by*) *experimentation*. Especially when uncertainty is large or when the time horizon is long, experimentation becomes more important (MacRae, 1972; Kendrick, 1982).[2] Later developments include multiple uncertainties (e.g., Wieland, 2000a), time-varying parameters (e.g., Beck and Wieland, 2002), alternative utility functions (e.g., Johnson, 2007), and correlated information (e.g., Marcoul and Weninger, 2008).[3]

In the Bayesian learning literature the decision maker has a prior belief on an uncertain parameter and she updates her belief with new information using Bayes' Rule. If the decision maker has no control over the precision of the new information or the rate at which it arrives, we refer to this as *passive learning*. For instance, Kelly and Kolstad (1999), Leach (2007), Kelly and Tan (2015), and Hwang et al. (2017) investigate the impact of passive learning about the climate system on optimal greenhouse gas emission control. In their models, learning is a by-product of control.

This paper differs from the literature. First and foremost, one of the control variables is used exclusively to increase the speed of learning – we refer to this as *active learning*. Specifically, we manipulate the precision of new observations, or the sharpness of the likelihood function in Bayes' Rule. In previous papers, learning occurs as time progresses (passive learning) or learning is a by-product of control (learning by

[2] In the literature, experimentation has been mainly investigated for monetary policy (e.g., Kendrick, 1982; Bertocchi and Spagat, 1993; Wieland 2000b; Beck and Wieland, 2002; Yetman, 2003).

[3] See Bar-Shalom and Tse (1976), Kendrick (1982; 2005) for non-Bayesian approaches.

experimentation). Because learning is a control variable, we have to take the cost of learning into account. Most other papers do not consider the cost of learning explicitly.[4] The gains from learning are improved decisions, whereas the costs of learning are losses. By balancing the gains and losses, the rate of learning is determined, together with the optimal control rate for emissions.

The implementation of active learning in a climate economy model is worthwhile since public as well as private funders subsidize research activities to learn about climate processes. The rate of learning depends on such efforts. For instance, WMO and UNEP (2010) estimate that global annual expenditures on climate observations are about $4~6 billion. As a result, observational errors on temperatures have substantially decreased (Kennedy et al. 2011). In addition, our approach highlights the channel of learning. We cannot experiment with the climate system. For climate sensitivity, defined as the equilibrium global warming in response to a doubling of the atmospheric concentration of carbon dioxide, we get only one observation a year (Knutti et al., 2002) from a natural, uncontrolled experiment. Considering the fact that the quality of the data can be improved when the observation is based on a more accurate measurement, our channel of learning (i.e., improving the quality of monitoring) is a useful way to increase the speed of learning, despite the limitations of the data.

This paper proceeds as follows. Section 2 describes the model and methods. The revised DICE model (Nordhaus, 2008) is used in this paper. Main revisions include the introduction of a fat-tailed distribution of climate sensitivity and the implementation of active learning about the climate sensitivity. Main results and sensitivity analyses are given in Section 3. Section 4 provides conclusions.

[4] Strictly speaking, the cost of learning is implicit in their papers in that experimentation with higher carbon emissions may lead to higher damages in the future. The magnitude of this effect has not been formally investigated yet.

2 The model and methods

2.1 The revised DICE model

We revise the well-known DICE model (Nordhaus, 2008) to introduce uncertainty and learning about climate change. Here we focus only on the main revisions. For the full model see Appendix A. Unless otherwise noted, initial values for the state variables and parameter values follow Nordhaus (2008). The decision maker in our model chooses the rate of emissions control and the level of research investment for each time period so as to maximize social welfare defined as in Equation (1). Gross output, net of damage costs and abatement costs, is allocated to consumption, research investment, and gross investment (other than climate research) as in Equation (2).

$$\max_{\mu_t, R_t} \mathbb{E} \sum_{t=0}^{\infty} L_t \beta^t \, U(C_t, L_t) \tag{1}$$

$$C_t = \left(1 - \theta_1 \mu_t^{\theta_2}\right) \Omega_t Q_t - I_t - R_t \tag{2}$$

where $\mathbb{E}$ is the expectation operator, U is the utility function, C is consumption, L is labor force, μ is the rate of emissions control, I is gross investment (other than climate research), Ω is the damage function, Q is gross output, R is investment in climate research, β is the discount factor, θ_1 and θ_2 are parameters.

The research capital stock accumulates as follows:

$$K_{R,t+1} = (1 - \delta_R) K_{R,t} + R_t \tag{3}$$

where K_R is the research capital stock, δ_R=0.15 is the depreciation rate of the research capital stock (Hall et al., 2010).[5]

Unlike DICE, we apply annual time steps with an infinite time horizon. In order to consider the effect of uncertainty more properly, we set the lower bound of consumption as low as possible and remove the upper bound of temperature increases. Reducing the computational burden, the savings rate is assumed to be exogenous ($I = sQ$).[6] As shown in Hwang et al. (2017), this assumption does not materially affect the results of DICE and its variants, since the savings rate does not change much for many plausible model specifications. In addition, backstop technology is not considered in the model. With backstop technology climate risk (e.g., catastrophic consumption loss) can be more easily eliminated by setting emissions to be zero at a finite cost.[7]

[5] δ_R=0.15 is generally assumed in the literature, although some recent studies find that δ_R is higher than 0.15 (Nadiri and Prucha, 1996; Hall et al., 2010; Li and Hall, 2016; Cristini et al., 2016). We find that the parameter value does not affect the main results of this paper from some sensitivity analyses (results not shown). The general finding is that research investment stabilizes at a low level, which varies according to δ_R, so as to compensate for the depreciated research capital stock.

[6] The computational method of this paper is based on Hwang (2017). The method obtains solutions from optimality conditions, and therefore we need to derive optimal policy rules from the first order conditions. This becomes a problem when a model has many control and state variables. Since our model has several nonlinear functions with 3 control variables and 10 endogenous state variables, deriving optimal policy rules is not an easy task. In addition, endogenous savings rate increases the computational burden. For a passive learning model of Hwang et al. (2017), even a single run (with a maximum tolerance level of 1e-4) was not solved within an hour (with a system of 2.5GHz Intel Core i7, 16GB 1600MHz DDR3), which means it would take several weeks for 1,000 Monte Carlo simulations. Regarding the constant savings rate, Hwang (2017) compares the results of the two versions of the DICE model (one with a constant savings rate, another with a flexible savings rate) and finds that there is no significant difference in the results between the two versions. One of the reasons is that the optimal savings rate of the DICE model does not change much over time and it gradually approaches a constant value as the economy approaches the equilibrium. For further discussions and accuracy tests see Hwang (2017). We defer the application of a flexible savings rate to future research.

[7] As Hwang et al. (2013; 2016) show, the application of fat-tailed distribution of the climate sensitivity with CRRA utility function into the DICE model (as in our model) may lead to a catastrophic consumption loss. It would be optimal for the decision maker to reduce carbon emissions totally (i.e. zero emission) if a backstop technology is available with a finite cost. This is an implication of the Dismal Theorem (Weitzman, 2009). The backstop technology in the original DICE model comes at a finite cost.

The temperature response model of the DICE model is:

$$T_{AT_{t+1}} = T_{AT_t} + \xi_1\{RF_{t+1} - (\eta/\lambda)T_{AT_t} - \xi_3(T_{AT_t} - T_{LO_t})\} \qquad (4)$$

$$T_{LO_{t+1}} = T_{LO_t} + \xi_4\{T_{AT_t} - T_{LO_t}\} \qquad (5)$$

where T_{AT} and T_{LO} are atmospheric and oceanic temperature changes, respectively, from 1900, $RF_{t+1} = \eta\ \ln(M_{AT,t}/M_b)/\ln(2) + RF_{N,t}$ is radiative forcing, M_{AT} is the carbon stock in the atmosphere, M_b is the pre-industrial carbon stock in the atmosphere, $RF_{N,t}$ is radiative forcing from other than greenhouse gas, λ is the equilibrium climate sensitivity, η, ξ_1, ξ_3, and ξ_4 are parameters.

This paper follows the tradition of Bayesian statistical decision theory which requires that uncertainty or partial ignorance can be represented as a probability distribution (DeGroot, 1970). We apply the framework of feedback analysis (Hansen et al., 1984) in order to introduce a fat-tailed climate sensitivity distribution.[8] Risk is fat-tailed if the probability density of an uncertain variable falls more slowly than exponentially in the tail (Weitzman, 2009). The probability distribution of the climate sensitivity is derived from the distribution of the total feedback factors (Roe and Baker, 2007), as

$$\lambda = \lambda_0/(1 - f) \qquad (6)$$

[8] The framework of feedback analysis is useful in the following reasons: 1) the total feedback factors are observable, unlike the climate sensitivity; 2) it is easy to apply the Bayes' Theorem since the total feedback factors are usually assumed to be normally distributed (Roe and Baker, 2007); 3) the resulting climate sensitivity distribution has fat tails.

where f is the total feedback factors, which is assumed to be strictly less than 1, and λ_0 is the equilibrium climate sensitivity in a black body planet without any feedbacks.[9] The total feedback factors denote the aggregate impacts of physical factors such as water vapor, cloud, and albedo on radiative forcing in a way to magnifying the response of the climate system (Stocker et al., 2013).

Substituting Equation (6) for λ in Equation (4), replacing radiative forcing with its components, rearranging, and introducing natural variability (stochastic shocks) result in:

$$T_{AT_{t+1}} = (\zeta_1 f + \zeta_2) T_{AT_t} + \zeta_3 \ln(M_{AT_t}/M_b) + \zeta_4 T_{LO_t} + \xi_1 RF_{N,t} + \varepsilon_{t+1}^{nv} \qquad (7)$$

where ε^{nv} is natural variability normally distributed with mean zero (Brohan et al., 2006), ζ_i (i=1,2,3,4) are adjusted parameters ($\zeta_1 = \xi_1 \eta / \lambda_0$, $\zeta_2 = 1 - \zeta_1 - \zeta_4$, $\zeta_3 = \xi_1 \eta / \ln(2)$, and $\zeta_4 = \xi_1 \xi_3$).

Observed atmospheric temperature is governed by the following equation.

$$T_{AT_{t+1}}^{obs} = T_{AT_{t+1}} + \varepsilon_{t+1}^{obs} \qquad (8)$$

where $T_{AT_t}^{obs}$ is the observed temperature change, ε^{obs} is observational errors normally distributed with mean zero including measurement errors and data coverage bias (Webster et al., 2008).

Notice that the climate system is not affected by observed temperature but by actual temperature. Actual temperature is not known to the decision maker and thus is not used for learning. Observed temperature affects the decision maker's belief on the climate sensitivity, resulting in the choice for the rate of emissions

[9] f should be strictly bounded above *on a physical science basis*. We know that λ is not defined for f=1 from Equation 6 and the climate system cannot be reached an equilibrium if f>1 (Baker and Roe, 2009).

control (or carbon tax). The climate system is indirectly affected by observed temperature in this way. We simulate both actual temperature and observed temperature and apply each for its own purpose: actual temperature for the climate system and observed temperature for the learning process.[10]

2.2 Implementation of active learning

The literature on decision making under uncertainty and learning about the climate sensitivity assumes that knowledge grows by one observation per year with constant precision (or observational error) (e.g., Kelly and Kolstad, 1999; Leach, 2007; Webster et al., 2008; Kelly and Tan, 2015; Hwang et al., 2017). Instead, this paper considers additional learning through improved monitoring, retaining annual observations. Research investment in the global climate observational system increases the precision not so much of individual temperature observations, but rather in the estimate of the global mean temperature. This in turn lowers estimation errors for the climate sensitivity (see Equation 12 below). As the standard error in the mean of temperature increases (decreases), the signal to noise ratio falls (grows), making it more difficult (easier) to detect the true state of the world.

---

[10] $T_{AT_t}^{obs}$ is a function of T_{AT_t} and observational errors, which is changing according to the level of research capital (see Equations 8, 11). Actual temperature is determined by the law of physics and natural variability, which do not change over time. Although actual temperature is not known to the decision maker, it affects future temperature as in Equation (7). The probability of high temperature increases at a certain point in time should be the same for both the passive learning model and the active learning model, if the choices of the decision maker are the same. But since the belief of the decision maker on the state of the world is different, her choices become different between the passive learning model and the active learning model. The change in expectation leads to the different decisions on the stringency of climate policy. If uncertainty is large (or fat-tailed), learning induces large changes in the choices of the decision maker.

As the global climate observational system improves, the standard error in the mean of temperature falls. For instance, as illustrated in Figure 1, the variance of global mean temperature has decreased over time as the number of weather stations has increased.[11]

[Figure 1]

From Equation (8), the standard error in the mean of observed temperature is decomposed into two elements as follows. We assume that the variance of natural variability is constant over time since natural variability is not controllable by the decision maker. In addition, we assume that natural variability and observational errors are independent.

$$\sigma_{\varepsilon,t}^2 = \sigma_{obs,t}^2 + \sigma_{nv}^2 \tag{9}$$

where σ_{ε}, σ_{ob}, and σ_{nv} are the standard error in the mean of observed temperature, observational errors, and natural variability, respectively.

Broadly speaking, observational errors are linearly related to the reciprocal of the number of observational instruments (Jones et al., 1997; Brohan et al., 2006), at least in the relevant domain (see Figure 1).[12] Assuming independence between sea surface temperature (SST) observational errors and land air temperature (LAT) observational errors, the total observational errors of global mean temperature can be

[11] The quality of observations as well as the number of observations is important for the standard error in the mean of temperature. For instance, increasing the number of observational system does not necessarily improve the precision of the measurement of climate change if the quality of observations is limited. However, for simplicity we refer the consideration of the quality of observations to future researches.

[12] The other non-linear relationships between uncertainty and research investment would be interesting to study, which is referred to future research.

calculated as follows. For simplicity, we assume that observational errors approach zero as investment in the global temperature observational system becomes arbitrarily large.

$$\sigma^2_{obs_t} = \omega_l \sigma^2_{obs_{l,t}} + \omega_s \sigma^2_{obs_{s,t}} = \sum_j \omega_j \left(\alpha_j / N_{oi\,j,t}\right) \qquad (10)$$

where $j \in \{l, s\}$ refers to each observation (l for LAT and s for SST), ω is the respective area of the land or the sea ($\omega_s + \omega_l = 1$), N_{oi} is the number of observational instruments, α_j is a parameter.

Equation (10) leads to Equation (11), the channel through which research investment affects the uncertainty about temperature shocks.

$$\sigma^2_{\varepsilon,t} = \omega_l\, c_l \alpha_l / \left(p K_{R_t}\right) + \omega_S c_s \alpha_s / \left\{(1-p) K_{R_t}\right\} + \sigma^2_{nv} = a_R / K_{R_t} + \sigma^2_{nv} \qquad (11)$$

where K_R is the research capital stock for the global temperature observational system, $0 \le p \le 1$ is the proportion of money spent on land observations, $c_l \equiv p K_R / N_{0,l}$ and $c_s \equiv (1-p) K_R / N_{0,s}$ are the unit cost of LAT and SST observation, respectively, and $a_R \equiv \omega_l\, c_l \alpha_l / p + \omega_S c_s \alpha_s / (1-p)$.

The decision maker updates her beliefs on the total feedback factors using Bayes' Rule as follows (Cyert and DeGroot, 1974; Lemoine, 2010):

$$p\left(f \middle| T^{obs}_{AT}\right) \propto p\left(T^{obs}_{AT} \middle| f\right) \times p(f) \qquad (12)$$

where $p(f)$ is the prior distribution, $p\left(T^{obs}_{AT} \middle| f\right)$ is the likelihood function (the likelihood of T^{obs}_{AT} observation given f), and $p\left(f \middle| T^{obs}_{AT}\right)$ is the posterior distribution. Since the temperature shocks are assumed to be normally distributed with mean zero, more precise temperature observations lead to an improved decision on carbon tax.

The normal distribution of the total feedback factors with parameters $\overline{f_t}$ and $\sigma_{f,t}$ is used as the initial prior for the year 2005 (Roe and Baker, 2007). The resulting posterior mean and the posterior variance of the total feedback factors are:

$$\overline{f_{t+1}} = \frac{\overline{f_t} + \zeta_1 T_{AT_t}^{obs} H_{t+1}\, v_t / v_{\varepsilon,t}}{1 + \left(\zeta_1 T_{AT_t}^{obs}\right)^2 v_t / v_{\varepsilon,t}} \tag{13}$$

$$v_{t+1} = \frac{v_t}{1 + \left(\zeta_1 T_{AT_t}^{obs}\right)^2 v_t / v_{\varepsilon,t}} \tag{14}$$

where $\overline{f_t}$ and $v_t = \sigma_{f,t}^2$ are the mean and the variance of the total feedback factors, $v_{\varepsilon,t} = \sigma_{\varepsilon,t}^2$ is the variance of temperature shocks, and $H_{t+1} \equiv T_{AT_{t+1}}^{obs} - \zeta_2 T_{AT_t}^{obs} - \zeta_3 \ln(M_t/M_b) - \zeta_4 T_{LO_t} - \zeta_5 RF_{N,t}$.

The posterior distribution with parameters $\overline{f_{t+1}}$ and v_{t+1} of Equations (13) and (14) serves as the prior for the next time period. In this way the decision maker learns about the true value of the total feedback factors for each time period. Note that the parameters of the posterior distribution are affected by research investment through Equations (11) and (14) and the parameters become endogenous state variables. The higher research investment, the lower is the variance of the total feedback factors. From Equation (6) the probability distribution of the climate sensitivity can be derived from the probability distribution of the total feedback factors.

2.3 Calibration

In order to calibrate a_R and σ_{nv}^2 in Equation (11) we use global expenditures on temperature observations. We take a bottom-up approach. Global mean LAT is calculated from the records of each country's weather stations and global mean SST is calculated from the reports of observational platforms such as ships, drifting buoys, and moored buoys (Kennedy et al., 2011). With this in mind, we multiply the number of each observational platform and the unit cost of each instrument (see Table 1). Annual operational costs for temperature observational instruments are estimated to be about $450 million in 2005.[13] The total installation costs for all the existing instruments are about $500 million.[14] Second, σ_{nv}^2 is calculated as the difference between the total variance of temperature shocks (=0.10^2) estimated by Tol and De Vos (1998) and the variance of observational errors (=0.06^2) obtained from the HadCRUT4 dataset (Morice et al., 2012). Then we can postulate that the research capital stock (K_{R_0}= $950 million) results in the variance of temperature shocks ($\sigma_{\varepsilon,0}^2$=0.10^2) (Equation 11). So, a_R=$3.42 million.[15]

[Table 1]

[13] For comparison, the United States spent $140 million on *in-situ* climate observations in 2010 (submission of USA to UNFCCC/SBI 35). WMO and UNEP (2010) estimate that annual global expenditure on climate observations is about $4~6 billion. Douglas-Westwood (2006) estimates that the total costs of ocean observations are $402 million in 2005. Their estimates are not directly comparable to this paper, however, because their estimates include all kinds of observations besides temperature, such as precipitation, wind, ice, as well as satellite observations.

[14] This number is fairly small compared to the world economy. The initial value for the global capital stock (in 2005) is $137 trillion in the original DICE model. Thus the research investment in climate observations has a negligible effect on the growth path of the world economy. In this case, cost-benefit analysis rather than a complex dynamic stochastic model can be used to calculate the optimal research investment. Nevertheless, our model is useful for the investigation of the effect of learning on optimal policy simultaneously with research investment. In addition, the literature on learning about climate change generally applies dynamic stochastic learning models (e.g., Kelly and Kolstad, 1999; Kelly and Tan, 2015).

[15] We do not separate the operational costs and the installation costs for these calibrations, for simplicity. That is, we assume that the research capital stock is the sum of the operational costs and the installation costs. An alternative is to explicitly represent operational costs in the model, but this does not affect the main results of this paper (results not shown).

As shown in Figure 2, our parameterization implies that the variance of temperature shocks decreases (increases) as the research capital stock increases (decreases). If there is no change in the research capital stock, the variance of observational errors (in turn, the variance of temperature shocks) remains the same. [16]

[Figure 2]

We set the level of satisfaction of the decision maker about the level of learning, mainly for computational reasons. This assumption also reflects the point that the decision maker does not want not make an effort to reduce the variance of observational errors if she thinks there has been enough learning. We can think of this as a budget constraint. This puts an upper bound on the research investment (in turn, on the research capital stock).[17] More specifically, we set R_t=0 if $\sigma_{\varepsilon,t}^2 - \sigma_{nv}^2 < \omega_c$, where ω_c reflects the level of satisfaction of the decision maker on the level of learning. According to our parameterization (σ_{nv}^2=0.064, ω_c=10^{-6}), if the variance of the temperature shocks $\sigma_{\varepsilon,t}^2$ (always higher than σ_{nv}^2, by assumption) becomes close to (but never equals to) the variance of natural variability σ_{nv}^2 through research investment, say 0.064<$\sigma_{\varepsilon,t}^2$<0.064001 (less than 0.002 percent higher than σ_{nv}^2=0.064), the decision maker does not invest money for learning.

This upper bound is binding after a certain year. For instance, when we set ω_c=10^{-6}, the upper bound of the research investment is binding after 2045. The problem is that the solution time increases a lot if there is no such an upper bound. For instance, a single run without the upper bound takes more than 30 minutes.

[16] Our calibration may overestimate the benefit of the research investment because temperature observations are not just used for learning about the climate sensitivity, but also for other various activities including weather forecast, agriculture, and transportation (AATSE, 2006). For the decomposition of the cost of temperature observations according to the sector of use, we refer to future research.

[17] From Equations (3) and (11) this assumption serves as an upper bound of research investment ($R_t \leq a_R/\omega_c - (1-\delta_R)K_{R,t}$). For instance, with ω_c=10^{-6} and the above parameterizations (i.e., a_R=\$3.42 million, $K_{R,0}$=\$950 million, δ_R=0.15), the upper bound of the initial research investment is about \$3.42 trillion.

With the upper bound, however, 1,000 Monte Carlo simulations take about 6 hours. The results between the two cases are a bit different as shown in Figure B.3 (Appendix B).

2.4 Computational methods

In order to solve our large model, we apply the computational method of Hwang (2017), who refines the earlier work by Maliar and Maliar (2005) and Judd et al. (2011).[18] The problem is reformulated in a recursive way as:

$$W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) = \max_{\boldsymbol{c}_t}[U(\boldsymbol{s}_t, \boldsymbol{c}_t, \boldsymbol{\theta}_t) + \beta \mathbb{E}_t W(\boldsymbol{s}_{t+1}, \boldsymbol{\theta}_{t+1})] \tag{15}$$

$$W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) \approx \sum_{n=1}^{N} \psi(\boldsymbol{s}_t, \boldsymbol{\theta}_t; \boldsymbol{b}_{\boldsymbol{n}}) \tag{16}$$

where W is the value function starting from period t, $\boldsymbol{c}$ is the vector of control variables (μ, R), $\boldsymbol{s}$ is the vector of state variables (K, K_R, M_{AT}, M_U, M_L, T_{AT}^{obs}, T_{LO}, $\bar{f}$, v, L, A, σ), M_U and M_L are the carbon stocks in the upper ocean and the lower ocean, respectively, σ is the emissions-output ratio, $\boldsymbol{\theta}$ is the vector of uncertain variables (f, ε), ψ is the basis function, and $\boldsymbol{b}$ is the vector of coefficients for the basis function.

The solution algorithm is summarized as follows. First, approximate the value function with a flexible basis function. Second, derive the first order conditions for optimal policy rules. Third, choose an initial guess on the coefficients $\boldsymbol{b}$ of the basis function: $\boldsymbol{b}^{(0)}$. Fourth, simulate a time series of variables satisfying

[18] Hwang (2017) takes advantage of the fact that the first order conditions of DICE result in global solutions (Solak et al., 2015). One of the advantages of the simulation-based method of Hwang (2017) is that it reduces the computational burden compared to the grid-based methods in the literature because it searches for solutions on a set satisfying the first order conditions. See Hwang (2017) for more discussions on the solution method.

the first order conditions, transitional equations, and boundary conditions with the initial guess $\boldsymbol{b}^{(0)}$. [19] Fifth, calculate the left hand side and the right hand side of the Equation (15) using the simulated time series, and then find $\boldsymbol{b}$ that minimizes the difference between them: $\widehat{\boldsymbol{b}}$. [20] Sixth, update the initial guess $\boldsymbol{b}^{(0)}$ using a pre-specified updating rule: $\boldsymbol{b}^{(1)}$. Seventh, iterate the above process with the new guess $\boldsymbol{b}^{(1)}$ until the value function converges. [21]

Accounting for random realizations of the uncertain variables, the model is run 1,000 Monte Carlo simulations and the average of all simulations is presented. For additional results, see Appendix B. The true value of the total feedback factors is set at 0.6 (which corresponds to the equilibrium climate sensitivity 3°C/2xCO_2, the most likely value according to the current scientific knowledge) (Stocker et al., 2013) for the reference case. The other true values are considered in Section 3.4.

For simulations, the initial values for $\overline{f_t}$ and v_t are assumed to be 0.65 and 0.13^2, respectively, following the current scientific knowledge (Roe and Baker, 2007). The models are also simulated with different initial beliefs but the general implications of these simulations do not change (results not shown). Since the total feedback factors are bounded above, the posterior distribution is derived first with the conjugate normal

[19] The simulation length is set at 1,000 years. Longer horizons do not affect the main results of this paper.

[20] The Gauss-Hermite integration is applied for the expectation in Equation (15) with 10 integration nodes. Higher number of nodes does not affect the main results of this paper.

[21] The maximum tolerance level is set at 1.0×10^{-4}, which is small relative to net present welfare. Existing numerical models generally apply a similar tolerance level (e.g., Kelly and Kolstad, 1999; Cai et al., 2012; Lemoine and Traeger, 2014). The more stringent tolerance level gains a bit more accuracy at the expense of the speed of convergence. Actually, changing the tolerance level from 1.0×10^{-4} to 1.0×10^{-5} does not change our results significantly. When the stopping criterion holds, the relative difference of two consecutive policy functions (i.e., $\max\left[\left|\{\mu_t^{(p+1)}-\mu_t^{(p)}\}/\mu_t^{(p)}\right|, \left|\{R_t^{(p+1)}-R_t^{(p)}\}/R_t^{(p)}\right|\right]$, where p refers to the pth iteration) is 6.2×10^{-4}. An alternative is to compare two consecutive policy functions for stopping criterion (i.e., $\max\left[\left|\{\mu_t^{(p+1)}-\mu_t^{(p)}\}/\mu_t^{(p)}\right|, \left|\{R_t^{(p+1)}-R_t^{(p)}\}/R_t^{(p)}\right|\right] \le \omega$, where ω is the maximum tolerance level) and compare the relative difference of value functions. We find that this does not affect our main results. If the relative difference of two consecutive policy functions was set at 1.0×10^{-3} for the tolerance level, we get the same optimal solutions and the relative difference of two consecutive value functions is 1.0×10^{-4}. For more error analyses see Appendix C.

prior, and then an upper bound ($\bar{f}_t \leq$0.999) is set for simulations following Hwang et al. (2017). The upper bound corresponds to the climate sensitivity of 1,200°C/2xCO_2, which is far higher than any admitted values. Note that the most likely value for the equilibrium climate sensitivity is 3°C/2xCO_2 (Stocker et al., 2013). Hwang et al. (2017) find that higher upper bounds than this do not affect the main results. Since the atmospheric CO_2 concentration is expected to be doubled around 2100, roughly speaking, the climate sensitivity is a proxy for temperature increases around 2100.

3 Results

3.1 The rate of learning

Figure 3 shows the evolution of the climate sensitivity distribution. For comparison, the results of the model with passive learning are also presented. As expected, the mean parameter $\bar{f}$ converges to the pre-specified true value and the variance parameter ν approaches – but never reaches – zero over time. The rate of learning, measured as the reduction in the (simulated) coefficient of variation of the climate sensitivity (Webster et al., 2008; Hwang et al., 2017), is higher under active learning than under learning only from temperature observations. It takes 44 years for the coefficient of variation to be reduced by 50% for improved observations, whereas it takes 57 years in the passive learning model. For comparison, the learning time for 50% reduction in the coefficient of variation of the climate sensitivity is about 60~70 years in Webster et al. (2008) when the prior similar to the current paper is used (see Figure 10 of their paper).[22] This is because by construction, learning in this paper constitutes an additional way to produce information.

[22] The rates of learning in Kelly and Kolstad (1999), Leach (2007), and Kelly and Tan (2015) are not directly comparable to the current paper since they define learning differently from ours: learning takes place in their models

[Figure 3]

Table 2 shows the corresponding probability of high temperature increases. The probability density in the upper tail of the climate sensitivity distribution shrinks faster for active learning than for passive learning. Therefore the probability of high temperature increases is lower in the active learning model than in the passive learning model. For instance, the probability of temperature increases higher than 6°C in 2105 for active learning is more than 5 times lower than the one for passive learning case. This is the reason why there is a substantial need for research investment, and the reason why there is a difference in carbon tax between the active learning case and the passive learning case (Section 3.3).

[Table 2]

3.2 Research investment

The optimal level of investment in climate research is much higher than the current level. For instance, the initial level of investment in the global climate observational system is about $1.2 trillion.[23] These are the reasons. First, the benefit of learning is far greater than the cost of learning (Keller et al., 2007a, b; Baehr et al., 2008). For comparison, with a discrete uncertainty representation and exogenous learning, Peck and Teisberg (1993) estimate that the value of information on the climate sensitivity is $148 billion, and Nordhaus and Popp (1997) estimate that the value of information on the climate sensitivity is $6.9 ~ 11.7

when the mean of the uncertain variable becomes statistically close (e.g., at a significance level of 0.05) to the pre-specified true value.

[23] We conduct sensitivity analyses on the level of satisfaction of the decision maker on learning (ω_c). If the decision maker wants less precise observations (see Section 2.3), the amount of money spent on the global observational system should be decreased. For instance, if the decision maker sets a criterion that ω_c=10^{-5} (10^{-4}, respectively), instead of 10^{-6} in the reference case, the level of investment is $341 billion ($34 billion, resp.). For those cases, however, the boundary condition (or the budget constraint) binds from the first period.

billion.[24] Keller et al. (2007b) estimate that the value of information on the climate sensitivity is about $10 billion.[25] Keller et al. (2007a) estimate that the value of information associated with early detection of changes in the North Atlantic meridional overturning circulation (MOC) is in the order of tens of billions of dollars, which is far higher than the cost of MOC observation systems (tens of millions of dollars, see Baehr et al., 2007). Baker and Solak (2010) estimate that the optimal level of investment in energy technology is on the order of tens of billions of dollars.[26] Since the current paper deals with the fat-tailed distribution of the climate sensitivity it is not surprising that the benefit of learning is greater in our model than in the literature (Weitzman, 2009).

Second, the research investment is small compared to the world economy. It is less than 1% of the global capital stock (see footnote 14). Of course, no one country would agree to pay the cost but we are concerned with a global model. In addition, the total cost of climate change (damage cost + abatement cost) can be largely reduced (see Figure 4) when we account for (active) learning because the decision maker can rule out, to some extent, the possibility of tail events (extreme temperature increases) from learning, and thus she spends less money for emissions control. For instance, the net present value of the avoided cost (the total cost of the uncertainty case minus the total cost of the active learning case) for the next two centuries

[24] Peck and Teisberg (1993) introduce 3 states of the world: 1, 3, and 5°C/2xCO_2. The value of information in their model is calculated as the difference in (monetized) expected utility between instant learning and no-learning. If learning in 40 years is considered, the value of information is $24 billion. Nordhaus and Popp (1997) consider 5 discrete climate sensitivities: mean, $\pm$1 standard deviation, and $\pm$2 standard deviation. The value of information in their model is calculated as the difference in (monetized) expected utility between instant learning and learning in 50 years.

[25] They consider 3 climate sensitivities. They also find that if there is a temperature limit of 2.5°C the value of learning about the climate sensitivity is $800 billion.

[26] AATSE (2006) and Cristini et al. (2016) find that the benefit of ocean observations is far larger than the cost of observations, although they do not account for learning about the climate sensitivity.

is about $ 9.2 trillion (applying the discount rate of DICE), which is 2.1% higher than the net present value of the total research investment during the next two centuries.[27]

Third, we model R&D as an unrestricted flow. Future research should consider capacity and expansion limits on R&D.

After the large initial investment, as shown in Figure 4, the research investment decreases sharply. The research investment from the second year is largely due to the depreciation of the research capital stock.[28] The research capital stock reaches an upper bound around 2045, implying that the decision maker satisfies with the level of learning after 2045, and thus does not spend money for improved observations (see Section 2.3) except the compensation for the depreciated research capital stock. This reflects the point that early investment to reduce uncertainty is more beneficial because (1) it benefits a longer future and (2) knowledge saturates in our model specifications. Note that research investment is here mostly in equipment rather than specialist personnel, so that a rapid scaling-up and –down is feasible. These results imply that the optimal decision is to make uncertainty as low as possible since the cost of learning is much lower than the benefit of learning. Figure 4 also shows some sensitivity analyses on the cost of learning. The research capital stock increases as the cost of learning increases, as implied by Equation (11).

[Figure 4]

---

[27] For a more thorough discussion, we need to consider the opportunity cost of research investment and the general equilibrium effect. Since our model is not well suited for such an analysis, we refer to future research.

[28] About 90% of the new investment after the initial year is to compensate for the depreciated research capital stock (results not shown).

3.3 Carbon tax

It is well known that fat-tailed risk substantially increases the stringency of climate policy. Even in the theoretical framework of Weitzman (2009), learning does not matter because the marginal damage cost (or willingness to pay for a reduction of carbon emission) becomes arbitrarily large under fat-tailed risk. In this case, the fat tail is not thinned by learning. However, as argued in the literature (Costello et al., 2010; Millner, 2013; Horowitz and Lange, 2014; Hwang et al., 2013, 2016), in a realistic setting such as bounded radiative forcing, the marginal damage cost can be bounded. Kelly and Tan (2015) and Hwang et al. (2017) apply this framework and find that learning substantially reduces the effect of fat-tailed risk because learning is faster in the tail. Policy recommendations such as carbon tax are substantially different from the no-learning case. The current paper follows this framework but extends the approach: active learning. The optimal carbon tax is calculated as a Pigovian tax (Nordhaus, 2008).[29] As expected, the optimal carbon tax is highest for the uncertainty model and is lowest for the deterministic model (see Figure 5 and Table 3). Active learning substantially reduces the carbon tax because it takes away so much uncertainty.[30]

[Figure 5]

As mentioned in Section 2, there are limits to learning. The variance of natural variability remains fixed at 0.08^2 for the reference case. This serves as the lower bound for the variance of temperature shocks. The

[29] The carbon tax is calculated as follows: $carbon\ tax_t = -\tau\,(\partial W/\partial E_t)/(\partial W/\partial K_t)$, where W is social welfare defined as the discounted sum of population-weighted utility of consumption over the whole time period (Equation 1), E is carbon emission, K is the capital stock, and τ is a constant. In our specification, with a climate sensitivity that is known to the analyst but not to the decision maker, social welfare differs with the true value of the climate sensitivity. The effect of the marginal change in carbon emission (and also in the capital stock) on social welfare also varies with the true value of the climate sensitivity.

[30] Considering the amount of carbon emissions and the use of carbon tax as a reference value for public policy making (Nordhaus, 2017), a seemingly small difference in carbon tax rates between the cases in Table 3 is, de facto, a very important difference. Annual CO_2 emissions from the combustion of fossil fuels and cement production are about 9.5 GtC in 2011 (Stocker et al., 2013). Simply put, one-dollar difference in carbon tax rate ($1/tC) leads to a difference of about $10 billion a year.

sensitivity of the optimal carbon tax to this lower bound is shown in Figure 6. The optimal carbon tax decreases if the lower bound of temperature shocks decreases. The higher the magnitude of learning, the lower is the optimal carbon tax.

[Figure 6]

Table 3 presents the results when the damage function of Weitzman (2012), a highly reactive damage function, is applied.[31]

$$\Omega_t = 1/\left[1 + \pi_1 T_{AT,t} + \pi_2 T_{AT,t}^{\pi_2} + \pi_3 T_{AT,t}^{\pi_4}\right] \quad (17)$$

where π_1=0, π_2=0.0028388, π_3=0.0000050703, and π_4=6.754. For the DICE damage function, π_1=0, π_2=0.0028388, and π_3=π_4=0.

The effect of the fat tail on climate policy is substantial when the damage function of Weitzman (2012) is applied (see the results of the uncertainty model), but is greatly reduced when learning is introduced. Research investment further enhances the effect of learning (compare the cases where the true value of the climate sensitivity is 3°C/2xCO_2). This signifies the importance of learning, especially when the damage function is highly responsive to higher temperature increases. Because we can rule out, to some extent, the impact of tail events (say, temperature increases of 4.5°C or more) from learning, the intensity of climate policy can be largely reduced. The value of research investment is that the decision maker can effectively reduce the probability of substantially high temperature increases as shown in Table 2. Thereby active learning helps to avoid an unnecessarily large amount of expenditure on emission reductions. Notice that

[31] The difference between the two damage functions becomes significant when temperature increases are higher than 3°C or more (Hwang et al., 2013).

the optimal carbon tax is a measure of the intensity of climate policy in this paper.

[Table 3]

3.4 The true value of the climate sensitivity

In the real world, the true value of the climate sensitivity is not known with certainty. This subsection investigates how sensitive our main results to the different true values of the climate sensitivity. Figure 7 and Table 2 show some results. Roughly speaking, future temperature evolves according to the true value of the climate sensitivity (although not known to the decision maker) and the decisions made with regard to the states resulting from the current decisions (e.g., emissions control rate, research investment). Other things being equal, future temperature will be higher if the true value of the climate sensitivity is higher as shown in Figure 7. We also observe that the higher the true value of the climate sensitivity, the fatter is the right tail of the climate sensitivity distribution. Since the decision maker considers future damages for the current decisions (Bellman's principle of optimality) (Bellman and Dreyfus, 1962), the rate of emissions control or the optimal carbon tax should be higher (lower, respectively) for the higher (lower, resp.) true value cases.

[Figure 7]

Table 3 and Figure 8 show the cases where the true value of the climate sensitivity (λ) is different from the initial guess. The initial belief on λ is the same as the reference case. Other things being equal, as expected, the optimal carbon tax is higher if the true λ is higher. The optimal carbon tax for the active learning model is lower than the one for the passive learning model, as in the reference case. Applying the damage function of Weitzman, the effect of active learning becomes more prominent. [32]

[32] These results may be sensitive to the model specifications including the probability distribution, the true value of the climate sensitivity, damage function, preference parameters, etc. Further analysis is deferred to future research.

[Figure 8]

4 Conclusions

We define active learning as investment in knowledge acquisition. Active learning is contrasted to passive learning where new information arrives without cost at a fixed rate and precision. We investigate the effect of active learning on greenhouse gas emission reduction, noting that climate policy is typified by (deep) uncertainty and irreversibility. The decision maker chooses to reduce uncertainty through significant investment in climate research, more than three orders of magnitude greater than the current level of expenditure. This helps the social planner make better decisions on climate policy. The level of uncertainty decreases more rapidly with active learning than with passive learning. As a result, the optimal carbon tax is lower for active learning than for passive learning, which in turn is lower than the carbon tax without learning. The effect of learning is more pronounced as the tail risk increases. Earlier investment in climate research is more beneficial than later investment.

This paper is the first to introduce active learning into an integrated assessment model of climate and the economy. Applying alternative ways of learning would help to understand the role of learning further including reconstructions of past temperature which would increase the number of observations in the likelihood, and investment in climate research which would sharpen the prior.

This paper applies the solution method demonstrated in Hwang (2017). His method allows us to solve multiple iterations of the model with many state variables quickly for the Monte Carlo analysis. Hwang (2017) suggests that this should come at little cost of model accuracy. If it does reduce accuracy, the likely impacts on the results are higher savings rates as implied by Hwang (2017). However, this issue is minor in our model because we apply a constant savings rate. As shown by Hwang (2017), the difference of the results obtained from the two solution methods is minimal when a constant savings rate is applied. Even if that were the case, our paper demonstrates new methodology that can be applied to richer models as computer resources improve.

Active learning apart, this paper closely follows Nordhaus' DICE model. Other specifications should be explored, including alternative utility functions (e.g., Sterner and Persson, 2008), different mitigation cost and climate impact functions, and other economic growth models. Perhaps more importantly, we here posit a true value of the climate sensitivity, rather than a probability density function, which exaggerates the effects of learning. We only consider one uncertain parameter, which suppresses both the effect of uncertainty and the cost of learning. We study the case of a global planner. Kolstad and Ulph (2008; 2011) show that uncertainty enhances cooperation. This would imply that active learning fosters free-riding. All these matters are deferred to future research.

Active learning by improved monitoring also applies to other areas of public policy. Learning by experimentation with policy variables is informative for issues with a short characteristic life time – monetary policy, for instance – but less so for issues that span long periods – besides climate change, pensions, education and structural unemployment come to mind. The method proposed here applies to any area in which knowledge of the response to policy is imperfect partly due to imperfect monitoring. Although not always acknowledged, most economic quantities are observed imperfectly or imprecisely, including population size, quality of education, occupation, prices, international trade and income. The importance of active learning through improved monitoring is therefore probably not limited to climate policy.

Acknowledgments

Appendix A: The full model

The list of variables and parameters are given in Tables A.1 and A.2.

$$\max_{\mu_t,R_t} \mathbb{E} \sum_{t=0}^{\infty} L_t \beta^t \, U(C_t, L_t) \tag{A.1}$$

$$C_t = \left(1 - \theta_1 \mu_t^{\theta_2}\right) \Omega_t Q_t - I_t - R_t \tag{A.2}$$

$$K_{R,t+1} = (1 - \delta_R) K_{R,t} + R_t \tag{A.3}$$

$$K_{t+1} = (1 - \delta_k) K_t + I_t \tag{A.4}$$

$$M_{AT\,t+1} = (1 - \mu_t)\sigma_t Q_t + E_{LAND_t} + \delta_{AA} M_{AT\,t} + \delta_{UA} M_{U\,t} \tag{A.5}$$

$$M_{U\,t+1} = \delta_{AU} M_{AT\,t} + \delta_{UU} M_{U\,t} \tag{A.6}$$

$$M_{L\,t+1} = \delta_{UL} M_{U\,t} + \delta_{LL} M_{L\,t} \tag{A.7}$$

$$\begin{aligned} T_{AT_{t+1}} = T_{AT_t} + \xi_1 \{ \eta \, \ln(M_t/M_b)/\ln(2) + RF_{N,t} - \eta T_{AT_t}/\lambda - \xi_3 (T_{AT_t} \\ - T_{LO_t}) \} + \varepsilon_{t+1}^{nv} \end{aligned} \tag{A.8}$$

$$T_{AT_{t+1}}^{obs} = T_{AT_{t+1}} + \varepsilon_t^{obs} \tag{A.9}$$

$$T_{LO_{t+1}} = T_{LO_t} + \xi_4 \{ T_{AT_t} - T_{LO_t} \} \tag{A.10}$$

$$\overline{f_{t+1}} = \frac{\overline{f_t} + \zeta_1 T_{AT_t}^{obs} H_{t+1} \, v_t / v_{\varepsilon,t}}{1 + \zeta_1^2 T_{AT_t}^{obs^2} \, v_t / v_{\varepsilon,t}} \tag{A.11}$$

$$v_{t+1} = \frac{v_t}{1 + \zeta_1^2 T_{AT_t}^{obs^2} \, v_t / v_{\varepsilon,t}} \tag{A.12}$$

$$v_{\varepsilon,t} = \alpha_{R_1} / K_{R_t} + \sigma_{nv}^2 \tag{A.13}$$

where $\mathbb{E}$ is the expectation operator given information at point in time t (annual).

[Table A.1], [Table A.2]

Appendix B: Additional results

Figure B.1 is the results for improved observations (the reference case in Section 3). Each figure is the average of 1,000 Monte Carlo simulations. This figure shows how each variable evolves over time. For instance the rate of emissions control gradually increases during the first 2~3 centuries and then reaches the value 1 (full abatement). The carbon stock gradually decreases after the rate of emissions control becomes equal to 1. Atmospheric temperature follows the same pattern with a time lag. The maximum temperature increases (from 1900) are less than 4°C (in the early 22nd century) for all the cases. There is an initial peak in research investment and then the level of research investment becomes trivial. Consumption and gross investment (other than research investment) grows continuously since our model is based on the DICE model, which represents continuous economic growth.

[Figure B.1]

Figure B.2 shows the results of all runs.

[Figure B.2]

Figure B.3 shows the research capital stock with and without the upper bound.

[Figure B.3]

Appendix C: Error Analysis

We compare social welfare obtained from our solution method and maximum social welfare obtained from grid-based simulations. Since there are 12 state variables, it is not easy to implement simulations using Chebyshev nodes for each state variable (e.g., 5 nodes for each state variable lead to more than 244 million nodes). Instead, we construct grid points based on control variables. We construct 1001 (equally spaced) grid points for the rate of emission control (from 0 to 1) and 51 (equally spaced) grid points for research investment (from 0.5 times the optimal level to 1.5 times the optimal level) each time period and calculate the corresponding social welfare (51,051 grid points in total each year). Then the maximum social welfare is calculated as in Equation C.1.

$$W_g(\boldsymbol{s}_t, \boldsymbol{\theta}_t) = \max_{\boldsymbol{c}_{t,i}} [U(\boldsymbol{s}_t, \boldsymbol{c}_{t,i}, \boldsymbol{\theta}_t) + \beta \mathbb{E}_t W(\boldsymbol{s}_{t+1,i}, \boldsymbol{\theta}_{t+1})] \qquad \text{(C.1)}$$

where i refers to grid points, W_g is the value function obtained from grid-based simulations, $\boldsymbol{c}$ is the vector of control variables, $\boldsymbol{s}$ is the vector of state variables, $\boldsymbol{\theta}$ is the vector of uncertain variables.

The error is defined as the maximum relative difference between the two levels of social welfare obtained from our solution method and from grid-based simulations.

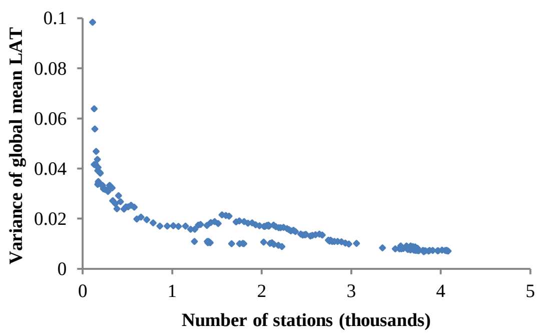
$$Error = \max_t \left[1 - W(\boldsymbol{s}_t, \boldsymbol{\theta}_t) / W_g(\boldsymbol{s}_t, \boldsymbol{\theta}_t) \right] \tag{C.2}$$

where W is the value function obtained from our solution method (Equation 15).

The error over the first 100 years is 1.6×10^{-4}. Figure C.1 shows some results. A wider domain may lead to different accuracy results, but this is not the case here because control variables are bounded in our specification (see Section 2.3). We do not find a wider domain (up to the upper bound of the variables) leads to less accuracy.

[Figure C.1]

Figures



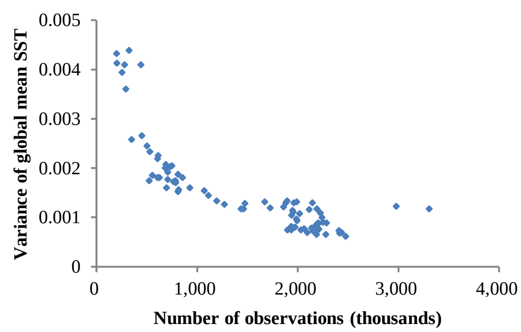


Figure 1 Uncertainty about global mean temperature (Left): The variance of global mean land air temperature (LAT) 1850-2006 (CRUTEM3, Brohan et al., 2006) as a function of the number of weather stations used to estimate the global mean temperature. **(Right):** The variance of global mean sea surface temperature (SST) 1925-2006 (HadSST3, Kennedy et al., 2011) as a function of the number of observations used to estimate the global mean temperature. The data were obtained from John Kennedy (personal communication).

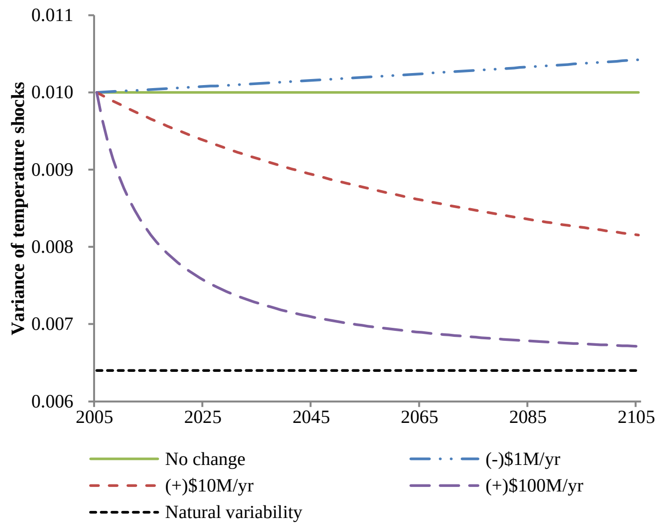


Figure 2 Hypothetical learning dynamics No change refers to the case where the research capita stock remains the same as in the initial year. + $X/yr (respectively, - $X/yr) refers to the case where the research capital stock increases (resp., decreases) $X every year from the initial level.

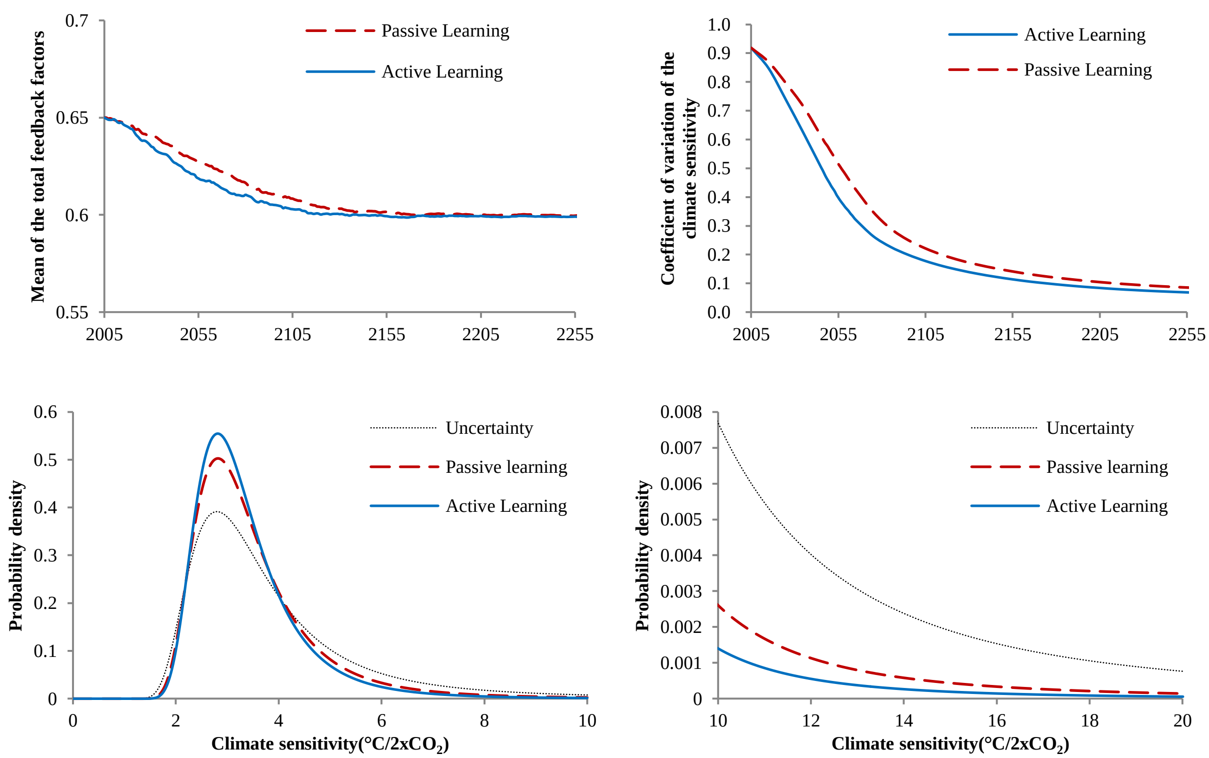


Figure 3 Climate sensitivity distribution (Top left): The mean of the total feedback factors **(Top right):** The (simulated) coefficient of variation of the climate sensitivity. **(Bottom left):** The climate sensitivity

distribution in 2055 (0~10°C/2xCO_2). **(Bottom right):** The climate sensitivity distribution in 2055 (10~20°C/2xCO_2).

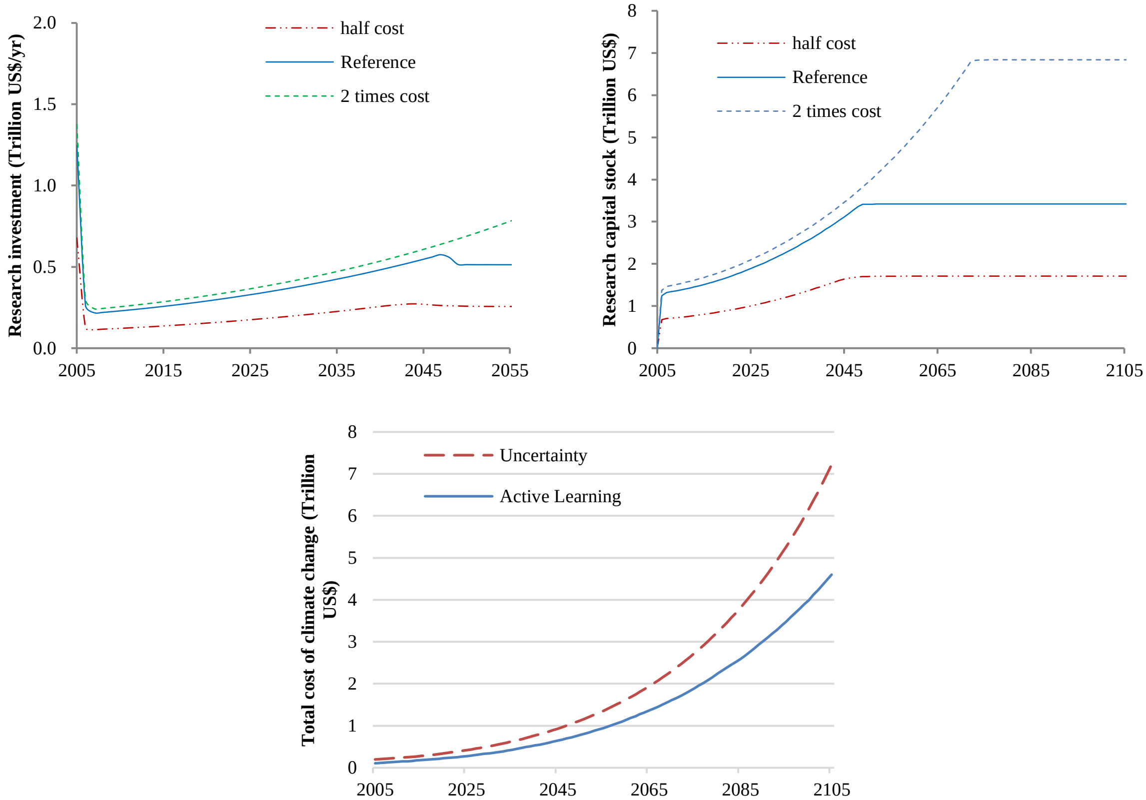


Figure 4 Research investment and the total cost of climate change (Top Left): The research investment **(Top Right):** The research capital stock **(Bottom):** The total cost of climate change (the gross output net of consumption, investment, and research investment) for the reference case and the uncertainty case.

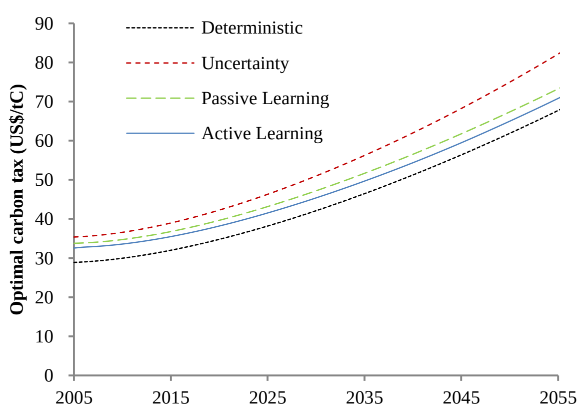


Figure 5 The optimal carbon tax

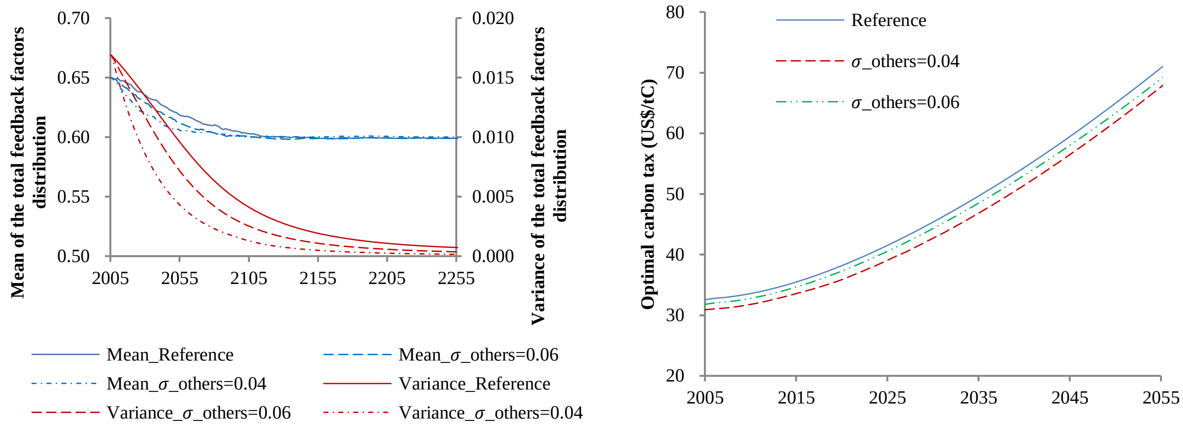


Figure 6 Limits to learning and the carbon tax (Left): The evolution of the learning parameters **(Right):** The optimal carbon tax.

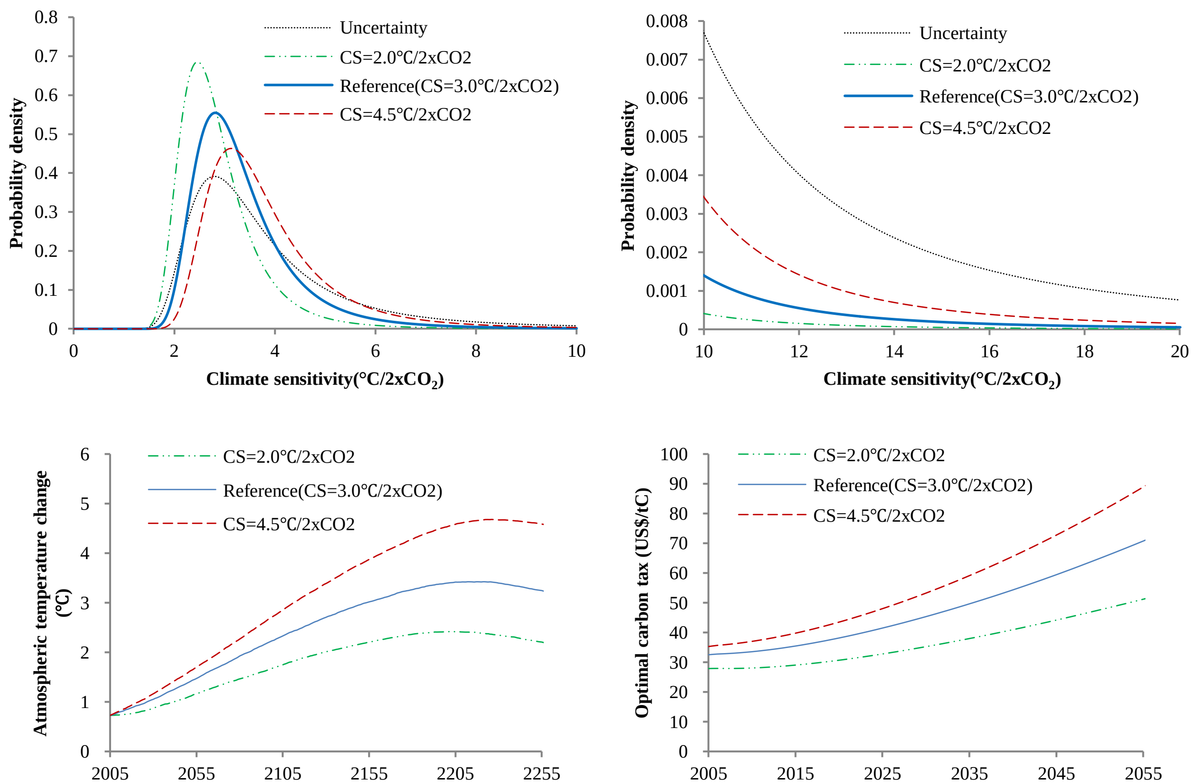


Figure 7 Sensitivity analysis (the true value of the climate sensitivity) CS refers to the climate sensitivity. **(Top panels):** The climate sensitivity distribution in 2055 **(Bottom left):** Atmospheric temperature **(Bottom right):** The optimal carbon tax

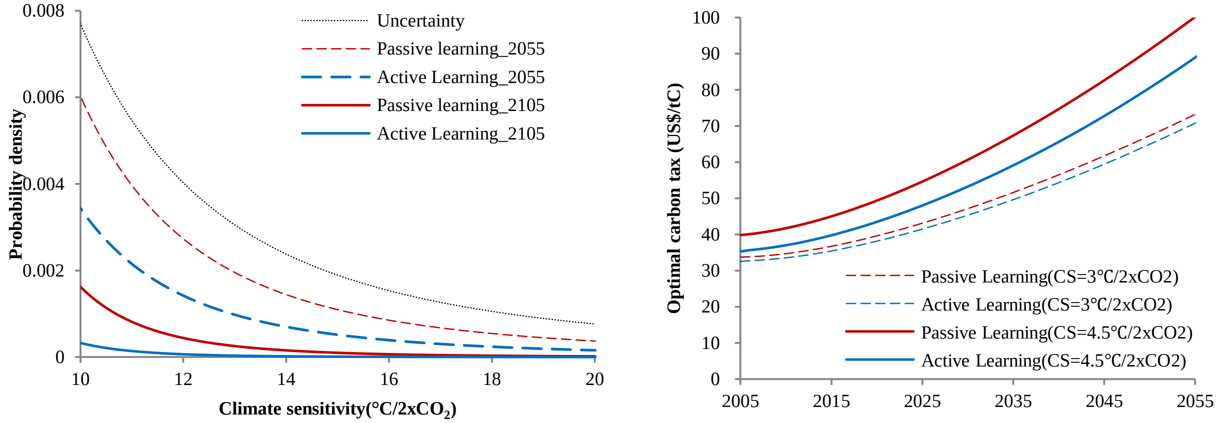


Figure 8 Sensitivity analysis (CS=4.5℃/2xCO_2) CS refers to the climate sensitivity **(Left):** The climate sensitivity distribution (the right segment higher than 10℃/2xCO_2) **(Right):** The optimal carbon tax

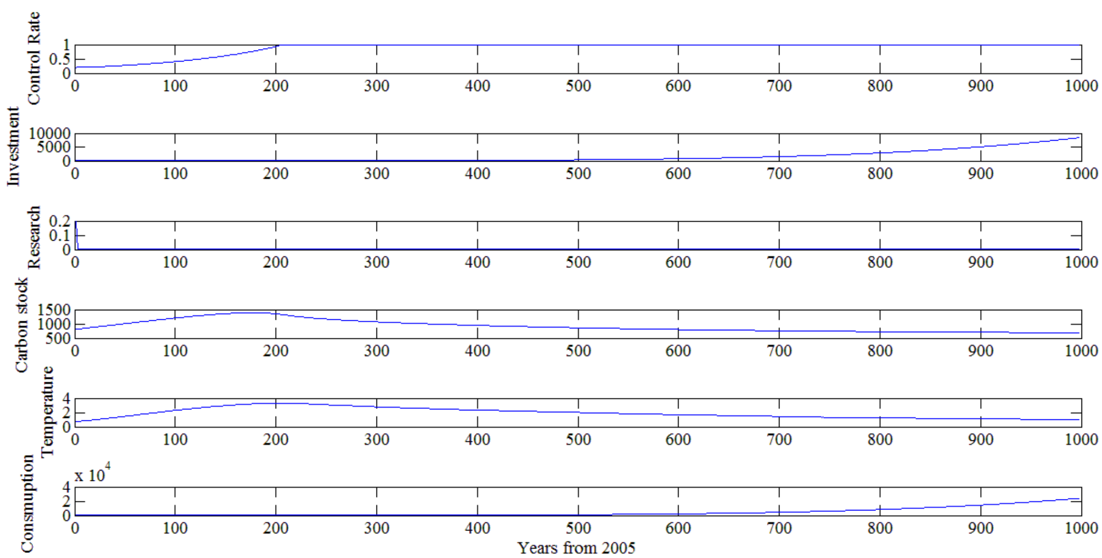


Figure B.1 Additional results The units for investment, research investment, the carbon stock, temperature increases, and consumption are $1,000 per person, trillion dollars, GtC, °C, and $1,000 per person, respectively.

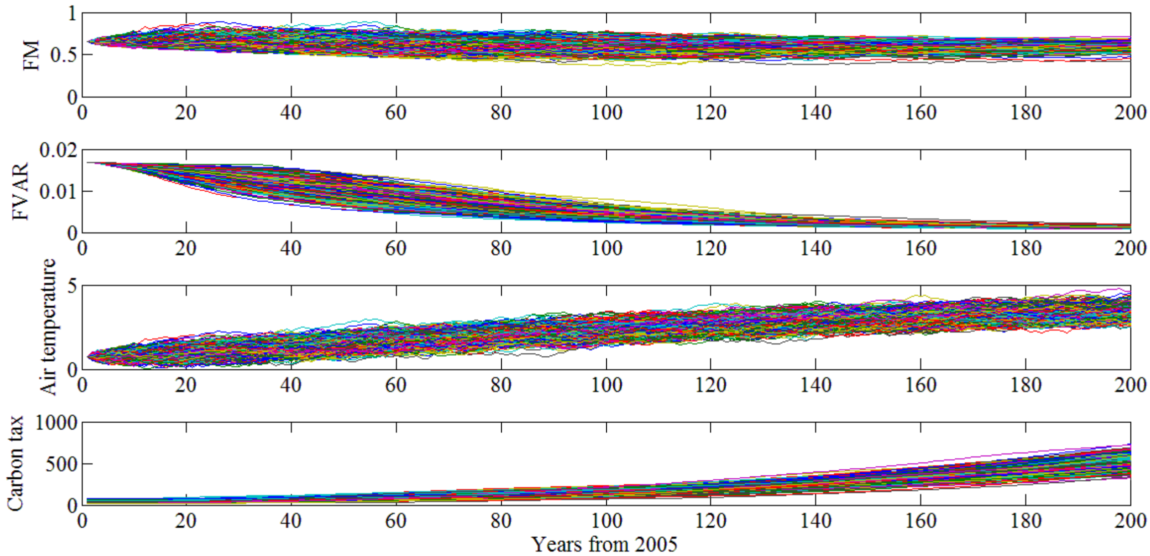


Figure B.2 Additional results (all simulations) (Top): The mean of the total feedback factors **(Upper middle):** The variance of the total feedback factors **(Lower middle):** Temperature increases (relative to 1900) **(Bottom):** The optimal carbon tax (US$/tC)

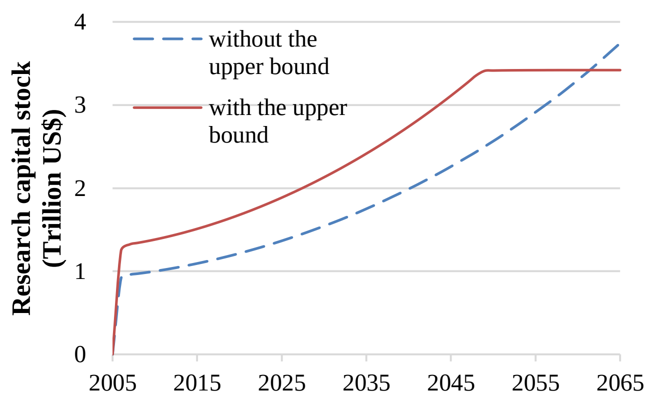


Figure B.3 Research capital stock with and without the upper bound Because of the computational burden, we present the average of 10 Monte Carlo simulations for the case with the upper bound (see Section 2.2). The initial research capital stock ($950 million) is so small that is close to zero in the figure.

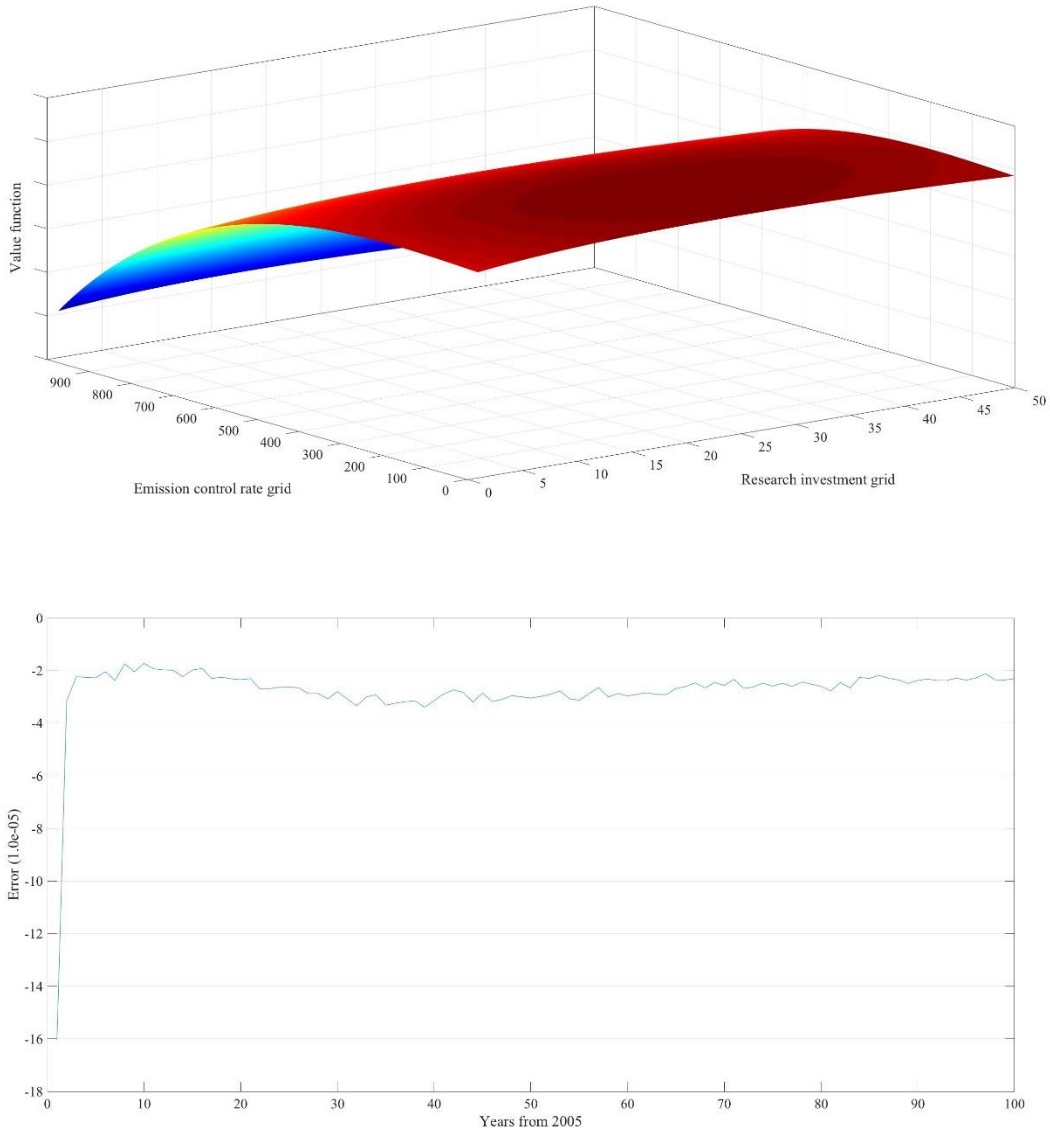


Figure C.1 Error analysis (Upper) Grid-based simulation results for the initial year. We do not present the value of z-axis since value function has a meaningless unit. **(Lower)** The relative difference of the value functions between the dynamic programming method and grid-based simulations for the initial 100 years.

Tables

Table 1 Global temperature observational system in 2005

	Number of instruments / observations (thousands)	Unit cost (1,000US$)			
		Installation		Operation (per year)	
		Low	High	Low	High

LAT	Weather station	3,455	40		60	
SST	Number of instruments					
	VOS	5,429			4	55
	Drifting Buoy	1,267			4.5	7.8
	Moored Buoy	194	1,150	2,700	200	500
	Number of Observations					
	VOS	1,169			0.00023	
	Drifting Buoy	1,632				
	Moored Buoy	179				
	Sum	2,980				

Note: The number of land weather stations is the one used for building the database CRUTEM4 (Kennedy et al., 2011). The number of voluntary observing ships, drifting buoys, and moored buoys are available at www.bom.gov.au/jcomm/vos and www.aoml.noaa.gov/phod/dac. The unit cost for land weather station is drawn from Mburu (2006). The unit costs for voluntary observing ships, drifting buoys, and moored buoys follow Kent et al. (2010), Meldrum et al., (2010), and Detrick et al., (2000), respectively. The unit cost for data transmission using satellite communication systems is about \$0.23 per observation (North, 2007). The number of SST observations is drawn from Kennedy et al. (2011).

Table 2 The probability of high temperature increase

	2055				2105			
	Passive learning, true λ=3	Active learning			Passive learning, true λ=3	Active learning		
		true λ=2	true λ=3	true λ=4.5		true λ=2	true λ=3	true λ=4.5
Prob. of λ>4.5	0.158	0.045	0.115	0.215	0.050	3.511E-04	0.019	0.211
Prob. of λ>6	0.053	0.010	0.031	0.068	0.006	1.089E-05	0.001	0.025
Prob. of λ>10	0.009	0.001	0.004	0.010	1.665E-04	6.419E-08	7.766E-06	4.152E-03

Note: $\boldsymbol{\lambda}$ is the equilibrium climate sensitivity (unit: ℃/2xCO_2). For all simulations, $\mathbb{E}_0 \boldsymbol{f}$=0.65, $\boldsymbol{\sigma_f}$=0.13.

Table 3 The optimal carbon tax in 2015 (US\$/tC)

	Deterministic true λ=3	Uncertainty $\bar{f}$=0.65, σ_f=0.13, true λ=3	Passive learning $\mathbb{E}_0 f$=0.65, σ_f=0.13			Active learning $\mathbb{E}_0 f$=0.65, $\mathbb{E}_0 \sigma_f$=0.13		
			true λ=2	true λ=3	true λ=4.5	true λ=2	true λ=3	true λ=4.5
DICE damage function	32.0	39.0	31.1	37.6	45.0	29.1	35.5	39.8
Weitzman's damage function	37.7	201.2	43.9	56.4	215.1	33.9	53.2	99.1

Note: $\boldsymbol{\lambda}$ is the equilibrium climate sensitivity (unit: ℃/2xCO_2).

Table A.1 Variables

U	Utility function	$=(C_t/L_t)^{1-\alpha}/(1-\alpha)$
C_t	Consumption	$=\left(1-\theta_1 \mu_t^{\theta_2}\right)\Omega_t Q_t - I_t - R_{i,t}$
μ_t	Emissions control rate	Control variable
R_t	Investment in climate research	Control variable
K_t	Capital stock	K_0=\$137 trillion
$K_{R,t}$	Research capital stock	$K_{R,0}$=\$950 million
$M_{AT\,t}$	Carbon stocks in the atmosphere	$M_{AT\,0}$=808.9GtC
$M_{U\,t}$	Carbon stocks in the upper ocean	$M_{U\,0}$=18,365GtC
$M_{L\,t}$	Carbon stocks in the lower ocean	$M_{L\,0}$=1,255GtC
T_{AT_t}	Atmospheric temperature deviations	T_{AT_0}=0.7307°C
T_{LO_t}	Ocean temperature deviations	T_{LO_0}=0.0068°C
$\bar{f}_t$	Mean of the total feedback factors	$\bar{f}_0$=0.65
v_t	Variance of the total feedback factors	$v_0=0.13^2$
Ω_t	Damage function	$=1/(1+\kappa_1 T_{AT_t}+\kappa_2 T_{AT_t}^{\kappa_3}+\kappa_4 T_{AT_t}^{\kappa_5})$
Q_t	Gross output	$=A_t K_t^{\gamma} L_t^{1-\gamma}$
I_t	Investment in general	$=sQ_t\Omega_t$
A_t	Total factor productivity	Exogenous
L_t	Labor force	Exogenous
σ_t	Emission-output ratio	Exogenous
$RF_{N,t}$	Radiative forcing from non-CO_2 gases	Exogenous
E_{LAND_t}	GHG emissions from the sources other than energy consumption	Exogenous
ε_t	Temperature shocks	Stochastic
$v_{\varepsilon,t}$	Variance of observed temperature shocks	$v_{\varepsilon,0}=0.1^2$

Note: The initial values for the state variables and the evolutions of the exogenous variables are from Cai et al. (2012), except for the research capital stock. The initial research capital stock does not affect the main results of this paper unless it is far higher than the default values. The lower bounds of the economic variables such as consumption, the capital stock, and gross world output are set to \$0.001 per person per year in this paper. In addition, there are no upper bounds for temperature increases.

Table A.2 Parameters

λ	Equilibrium climate sensitivity	$=\lambda_0/(1-f)$
f	True value of the total feedback factors	0.6
λ_0	Reference climate sensitivity	1.2°C/2xCO_2
s	Savings rate	0.245
α	Elasticity of marginal utility	2
ρ	Pure rate of time preference	0.015
β	Discount factor	$=1/(1+\rho)$
γ	Elasticity of output with respect to capital	0.3
δ_k	Depreciation rate of the capital stock	0.1
δ_R	Depreciation rate of research investment	0.15
κ_1, κ_2, κ_3, κ_4	Damage function parameters	κ_1=0, κ_2=0.0028388, κ_3=2, $\kappa_4=\kappa_5$=0
θ_1, θ_2	Abatement cost function parameters	θ_1=0.0561, θ_2=2.887
δ_{AA}, δ_{UA}, δ_{AU}, δ_{UU}, δ_{UL}, δ_{LL}, ξ_1, ξ_3, ξ_4, η	Climate parameters	δ_{AA}=0.9810712, δ_{UA}=0.0189288, δ_{AU}=0.0097213, δ_{UU}=0.005, δ_{UL}=0.0003119, δ_{LL}=0.9996881, ξ_1=0.022, ξ_3=0.3, ξ_4=0.005, η=3.8

α_R	Learning parameters	\$3.42 million
σ_{nv}^2	Learning parameters	0.0064
M_b	Pre-industrial carbon stock	596.4GtC
ω_c	Parameter reflecting satisfaction of the decision maker with the magnitude of learning	10^{-5}
$v_{\varepsilon,0}$	Initial value of the variance of observed temperature shocks	0.10^2

Note: The parameter values for climate parameters are from Cai et al. (2012). The parameter values for λ_0 and δ_R are from Roe and Baker (2007), Hall et al. (2010), respectively. The other parameters are from Nordhaus (1994; 2008) except for the learning parameters.

References

AATSE, 2006. Economics of Australia's sustained ocean observation system, benefits and rationale for public funding. Australian Academy of Technological Sciences and Engineering.

Baehr, J., Keller, K., Marotzke, J., 2008. Detecting potential changes in the meridional overturning circulation at 26 N in the Atlantic, Climatic Change 91, 11-27.

Baker, M.B., Roe, G.H., 2009. The shape of things to come: why is climate change so predictable? Journal of Climate 22, 4574–4589.

Baker, E., Solak, S., 2010. Climate change and optimal energy technology R&D policy, European Journal of Operational Research, 213(2), 442-454.

Bar-Shalom, Y., Tse, E., 1976. Caution, probing, and the value of information in the control of uncertain systems. Annals of Economic and Social Measurement 5(3), 323-337.

Beck, G.W., Wieland, V., 2002. Learning and control in a changing economic environment, Journal of Economic Dynamics and Control 26, 1359-377.

Bellman, R., Dreyfus, S.E., 1962. Applied dynamic programming, The RAND Corporations.

Bertocchi, G., Spagat, M., 1993. Learning, experimentation, and monetary policy. Journal of Monetary Economics 32, 169-183.

Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F.B., Jones, P.D., 2006. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. Journal of Geophysical Research, 111, D12106.

Cai, Y., Judd, K.L., Lontzek, T.S., 2012. DSICE: A dynamic stochastic integrated model of climate and economy. The Center for Robust Decision Making on Climate and Energy Policy Working Paper No. 12-02.

Costello, C.J., Neubert, M.G., Polasky, S.A., Solow, A.R., 2010. Bounded uncertainty and climate change economics. Proceedings of the National Academy of Sciences 107(18), 8108-8110.

Cristini, L., Lampitt, R.S., Cardin, V., Delory, E., Haugan, P., O'Neil, N., Petihakis, G., Ruhl, H.A., 2016. Cost and value of multidisciplinary fixed-point ocean observatories. Marine Policy 71, 138-146.

Cyert, R.M., DeGroot, M.H., 1974. Rational expectations and Bayesian analysis. The Journal of Political Economy 82, 521-536.

DeGroot, M.H., 1970. Optimal statistical decisions. McGraw-Hill, Inc.

Detrick, R., Frye, D., Collins, J., Gobat, J., Grosenbaugh, M., Petitt, R., Plueddeman, A., der Heydt, K., Wooding, F.B., Orcutt, J., 2000. DEOS moored buoy observatory design study. Woods Hole Oceanographic Institution Technical Report, August.

Douglas-Westwod, 2006. Global markets for ocean observation systems, Final OOS report.

Fitzpatrick, L.G., Kelly, D.L., 2017. Probabilistic stabilization targets. Journal of the Association of Environmental and Resource Economists 4(2), 611-657.

Gollier, C., Jullien, B., Treich, N., 2000. Scientific progress and irreversibility: An economic interpretation of the precautionary principle. Journal of Public Economics 75**,** 229-253.

Grossman, S.J., Kihlstrom, R.E., Mirman, L.J., 1977. A Bayesian approach to the production of information and learning by doing. The Review of Economic Studies 44(3), 533-547.

Hall, B.H., Mairesse, J., Mohnen, P. 2010. Measuring the Returns to R&D. Handbook of the Economics of Innovation, 2, 1033-1082.

Hansen, J., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., Ruedy, R., Lerner, J., 1984. Climate sensitivity: Analysis of feedback mechanisms. Geophysical Monograph Series 29, 130-163.

Horowitz, J. Lange, A., 2014. Cost-benefit analysis under uncertainty: A note on Weitzman's dismal theorem. Energy Economics 42, 201-203.

Hwang, I.C., Reynès, F., Tol, R.S.J., 2013. Climate policy under fat-tailed risk: An application of DICE. Environmental and Resource Economics 56(3), 415-436.

Hwang, I.C., Tol, R.S.J., Hofkes, M., 2016. Fat-tailed risk about climate change and climate policy. Energy Policy 89, 25-35.

Hwang, I.C., Reynès, F., Tol, R.S.J., 2017. The effect of learning on climate policy under fat-tailed risk. Resource and Energy Economics 48, 1-18.

Hwang, I.C., 2017. A recursive method for solving a climate-economy model: Value function iterations with logarithmic approximations. Computational Economics 50(1), 95-110.

Jensen, S., Traeger, C.P., 2013. Optimally climate sensitive policy under uncertainty and learning. Department of Agricultural and Resource Economics, UC Berkeley.

Johnson, T.C., 2007. Optimal learning and new technology bubbles. Journal of Monetary Economics 54, 2486-2511.

Jones, P.D., Osborn, T.J., Briffa, K.R., 1997. Estimating sampling errors in large-scale temperature averages. Journal of Climate 10, 2548-2568.

Karp, L., Zhang, J., 2006. Regulation with anticipated learning about environmental damages. Journal of Environmental Economics and Management 51, 259-279.

Keller, K., Deutsch, C., Hall, M.G., Bradford, D.F., 2007a. Early detection of changes in the North Atlantic meridional overturning circulation: implications for the design of ocean observation systems. Journal of Climate 20, 145-157.

Keller, K., Kim, S.R., Baehr, J., Bradford, D.F., Oppenheimer, M., 2007b. What is the economic value of information about climate thresholds? Human-Induced Climate Change: an Interdisciplinary Assessment, Schlesinger,M., Kheshgi, H., Smith, J., De La Chesnaye, F., Reilly, J.M., Wilson, T., Kolstad, C. (eds).Cambridge University Press: Cambridge 343-354.

Kelly, D.L., Kolstad, C.D., 1999. Bayesian learning, growth, and pollution. Journal of Economic Dynamics and Control 23, 491-518.

Kelly, D.L., Tan, Z., 2015. Learning and Climate Feedbacks: Optimal Climate Insurance and Fat Tails. Journal of Environmental Economics and Management 72: 98-122.

Kendrick, D., 1982. Caution and probing in a macroeconomic model. Journal of Economic Dynamics and Control 4, 149-170.

Kendrick, D.A., 2005. Stochastic control for economic models: Past, present and the paths ahead. Journal of Economic Dynamics and Control 29, 3-30.

Kennedy, J.J., Rayner, N.A., Smith, R.O., Parker, D.E., Saunby, M., 2011. Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. Journal of Geophysical Research 116, D14104.

Kent, E., Hall, A.D., Leader VOSClim Task Team, 2000. The voluntary observing ship (VOS) scheme. American Geophysical Union, 2000 Florida Ave., N. W. Washington DC 20009 USA.

Kolstad, C.D., 1996a. Learning and stock effects in environmental regulation: The case of greenhouse gas emissions. Journal of Environmental Economics and Management 31, 1-18.

Kolstad, C.D., 1996b. Fundamental irreversibilities in stock externalities. Journal of Public Economics 60, 221-233.

Kolstad, C.D., Ulph, A., 2008 Learning and international environmental agreements. Climatic Change 89,125–141.

Kolstad, C.D., Ulph, A., 2011. Uncertainty, Learning, and Heterogeneity in International Environmental Agreements. Environmental Resources Economics 50, 289-403.

Knutti, R., Stocker, T.F., Joos, F., Plattner, G., 2002. Constraints on radiative forcing and future climate change from observations and climate model ensembles. Nature 46(18), 719-723.

Leach, A.J., 2007. The climate change learning curve. Journal of Economic Dynamics and Control 31, 1728-1752.

Lemoine, D.M., 2010. Climate sensitivity distributions depend on the possibility that models share biases, Journal of Climate 23(16), 4395-4415.

Lemoine, D.M., Traeger, C., 2014. Watch Your Step: Optimal Policy in a Tipping Climate, AEJ-Economic Policy 6(1), 137-166.

Li, W.C., Hall, B.H., 2016, Depreciation of business R&D capital (No. w22473). National Bureau of Economic Research.

MacRae, E.C., 1972. Linear decision with experimentation. Annals of Economic and Social Measurement 1(4), 437-447.

Manne, A., Richels, R.G., 1992. Buying greenhouse insurance: the economic costs of carbon dioxide emission limits. The MIT Press.

Marcoul, P., Weninger, Q., 2008. Search and active learning with correlated information: Empirical evidence from mid-Atlantic clam fishermen. Journal of Economic Dynamics and Control. 32, 1921-1948.

Mburu, D.N., 2006. Possibilities for expansion of surface weather observing systems in east Africa. WMO publications IOM-94-TECO2006 1(6).

Meldrum, D., Wallace, A., Rolland, J., Burnett, W., Lumpkin, R., Niller, P., Viola, H., Charpentier, E., Fedak, M., 2010. Data buoy observations: The status quo and anticipated developments over the next decade. Proceedings of OceanObs 9.

Millner, A., 2013. On welfare frameworks and catastrophic climate risks. Journal of Environmental Economics and Management 65, 310-325.

Morice, C.P., Kennedy, J.J., Rayner, N.A., Jones, P.D., 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. Journal of Geophysical Research 117, D08101.

Nadiri, M.I., Prucha, I.R., 1996. Estimation of the depreciation rate of physical and R&D capital in the U.S. total manufacturing sector. Economic Inquiry 34(1), 43-56.

Nordhaus, W.D., 1994. Managing the Global Commons: The Economics of Climate Change. The MIT Press.

Nordhaus, W.D., 2008. A question of balance: Weighing the options on global warming policies. New Haven and London: Yale University Press.

Nordhaus, W.D., 2017. Revisiting the social cost of carbon. Proceedings of National Academy of Science 114(7), 1518-1523.

Nordhaus, W.D., Popp, D., 1997. What is the value of scientific knowledge? An application to global warming using the PRICE model. The Energy Journal 18, 1-46.

North, S., 2007. Report of the task team on satellite communications costs. SOT-IV. 16-21 April 2007, Geneva, Switzerland.

Peck, S.C., Teisberg, T.J., 1993. Global warming uncertainties and value of information: An analysis using CETA. Resource and Energy Economics 15, 71-97.

Prescott, E.C., 1972. The multi-period control problem under uncertainty. Econometrica, 40(6), 1043-1058.

Roe, G.H., Baker, M.B., 2007. Why is climate sensitivity so unpredictable? Science 318(5850), 629-632.

Roe, G.H., 2009. Feedbacks, timescales, and seeing red. Annual Review of Earth and Planetary Sciences 37, 93-115.

Rudik, I., 2016. Optimal climate policy when damages are unknown. Iowa State University Working Ppaer.

Solak, S., Baker, E., Chen, H., 2015. Convexity analysis of the dynamic integrated model of climate and the economy (DICE). Environmental Modeling and Assessment 20, 443–451.

Sterner, T., Persson, U.M., 2008. An even sterner review: Introducing relative prices into the discounting debate. Review of Environmental Economics and Policy 2, 61-76.

Stocker, T.F., Qin, D., Platter, G., Tignor, M.M.B., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., 2013. Summary for policymakers. In Climate change 2013: The physical science basis. Working Group 1 Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Tinbergen, J., 1954. Centralization and decentralization in economic policy. Amsterdam, North-Holland.

Tol, R.S.J., 2003. Is the uncertainty about climate change too large for expected cost-benefit analysis? Climatic Change 56(3), 265-289.

Tol, R.S.J., De Vos, A.F., 1998. A Bayesian statistical analysis of the enhanced greenhouse effect. Climatic Change 38(1), 87-112.

Ulph, A., Ulph, D., 1997. Global warming, irreversibility and learning. The Economic Journal 107, 636-650.

Van Wijnbergen, S., Willems, T., 2015. Optimal learning on climate change: Why climate skeptics should reduce emissions. Journal of Environmental Economics and Management 70, 17-33.

Webster, M., 2002. The curious role of learning in climate policy: Should we wait for more data? The Energy Journal 23(2), 97-119.

Webster, M., Jakobovits, L., Norton, J., 2008. Learning about climate change and implications for near-term policy. Climatic Change 89, 67-85.

Weitzman, M.L., 2009. On modeling and interpreting the economics of catastrophic climate change. Review of Economics and Statistics 91(1), 1-19.

Weitzman, M.L., 2012. GHG targets as insurance against catastrophic climate damages. Journal of Public Economic Theory 14(2), 221-244.

Wieland, V., 2000a. Learning by doing and the value of optimal experimentation. Journal of Economic Dynamics and Control 24, 501-534.

Wieland, V., 2000b. Monetary policy, parameter uncertainty and optimal learning. Journal of Monetary Economics 46, 199-228.

WMO, UNEP, 2010. Implementation plan for the global observing system for climate in support of the UNFCCC (2010 update). GOOS-184 GTOS-76 WMO-TD No. 1523.

Yetman, J., 2003. Probing potential output: Monetary policy, credibility, and optimal learning under uncertainty. Journal of Macroeconomics 25, 311-330.